On a Ramsey-type problem of Erdős and Pach

Ross J. Kang[*] Eoin Long[†] Viresh Patel[‡] Guus Regts[§]

February 23, 2016

Abstract

In this paper we show that there exists a constant $C > 0$ such that for any graph G on $Ck \ln k$ vertices either G or its complement $\overline{G}$ has an induced subgraph on k vertices with minimum degree at least $\frac{1}{2}(k-1)$. This affirmatively answers a question of Erdős and Pach from 1983.

Keywords: Ramsey theory, quasi-Ramsey numbers, graph discrepancy, probabilistic method.
MSC: 05C55 (Primary) 05D10, 05D40 (Secondary).

1 Introduction

Recall that the (diagonal, two-colour) Ramsey number is defined to be the smallest integer $R(k)$ for which any graph on $R(k)$ vertices is guaranteed to contain a homogeneous set of order k — that is, a set of k vertices corresponding to either a complete or independent subgraph. The search for better bounds on $R(k)$, particularly asymptotic bounds as $k \to \infty$, is a challenging topic that has long played a central role in combinatorial mathematics (see [4, 7]).

We are interested in a degree-based generalisation of $R(k)$ where, rather than seeking a clique or coclique of order k, we seek instead an induced subgraph of order (at least) k with high minimum degree (clique-like graphs) or low maximum degree (coclique-like graphs). Erdős and Pach [1] introduced this class of problems in 1983 and called them *quasi-Ramsey problems*. By gradually relaxing the degree requirement, a spectrum of Ramsey-type problems arise, and Erdős and Pach showed that this spectrum exhibits a sharp change in behaviour at a certain point. Naturally, this point corresponds to a degree requirement of half the order of the subgraph sought. Three of the authors recently revisited this topic together with Pach [5], and refined our understanding of the threshold for mainly what is referred to in [5] as the *variable quasi-Ramsey numbers* (corresponding to the parenthetical 'at least' above). In the present paper we focus on the harder version of this problem, the

[*]Radboud University Nijmegen. Email: ross.kang@gmail.com.

[†]Tel Aviv University. Email: eoinlong@post.tau.ac.il

[‡]University of Amsterdam. Email: viresh.s.patel@gmail.com. Supported by the Queen Mary - Warwick Strategic Alliance and the Netherlands Organisation for Scientific Research (NWO) through the Gravitation Programme Networks (024.002.003).

[§]University of Amsterdam. Email: guusregts@gmail.com. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 339109. This author was supported by a NWO Veni grant.

determination of what is called the *fixed quasi-Ramsey numbers* (where 'exactly' is implicit instead of 'at least' above).

Using a result on graph discrepancy, Erdős and Pach [1] proved that there is a constant $C > 0$ such that for any graph G on at least Ck^2 vertices either G or its complement $\overline{G}$ has an induced subgraph on k vertices with minimum degree at least $\frac{1}{2}(k-1)$. With an unusual random graph construction, they also showed that the previous statement does not hold with $C'k \ln k / \ln \ln k$ in place of Ck^2 for some constant $C' > 0$. They asked if it holds instead with $Ck \ln k$. (This was motivated perhaps by the fact that this bound holds for the corresponding variable quasi-Ramsey numbers.) Our main contribution here is to confirm this, by showing the following.

Theorem 1. *There exists a constant $C > 0$ such that for any graph G on $Ck \ln k$ vertices, either G or its complement $\overline{G}$ has an induced subgraph on k vertices with minimum degree at least $\frac{1}{2}(k-1)$.*

Although it is short, our proof of Theorem 1 has a number of different ingredients, including the use of graph discrepancy in Section 2, an application of the celebrated 'six standard deviations' result of Spencer [8] in Section 3 and a greedy algorithm in Section 4 that was inspired by similar procedures for max-cut and min-bisection. It is interesting to remark that the two discrepancy results we use are of a different nature; the one in Section 2 is an anti-concentration result while the result of Spencer is a concentration result.

2 An auxiliary result via graph discrepancy

Our first step in proving Theorem 1 will be to apply the following result. This is a bound on a variable quasi-Ramsey number which is similar to Theorem 3(a) in [5]. The idea of the proof of this auxiliary result is inspired by the sketch argument for Theorem 2 in [1], in spite of the error contained in that sketch (cf. [5]).

Theorem 2. *For any constant $\nu \geq 0$, there exists a constant $C = C(\nu) > 1$ such that for any graph G on $Ck \ln k$ vertices, G or its complement $\overline{G}$ has an induced subgraph on $\ell \geq k$ vertices with minimum degree at least $\frac{1}{2}(\ell-1) + \nu\sqrt{\ell-1}$.*

Note that the $O(k \ln k)$ quantity is tight up to an $O(\ln \ln k)$ factor by the unusual construction in [1] (cf. also Theorem 4 in [5]). The astute reader may later notice that the second-order term $\nu\sqrt{\ell-1}$ in the minimum degree guarantee of Theorem 2 can be straightforwardly improved to an $\Omega(\sqrt{(\ell-1)\ln \ln \ell})$ term. Since this does not seem to help in our results, we have omitted this improvement to minimise technicalities. On the other hand, a standard random graph construction yields the following, which certifies that the second-order term cannot be improved to a $\omega(\sqrt{(\ell-1)\ln \ln \ell})$ term.

Proposition 3. *For any $c > 0$, for large enough k there is a graph G with at least $k \ln^c k$ vertices such that the following holds. If H is any induced subgraph of G or $\overline{G}$ on $\ell \geq k$ vertices, then H has minimum degree less than $\frac{1}{2}(\ell-1) + \sqrt{3c(\ell-1)\ln \ln \ell}$.*

Proof. Substitute $\nu(\ell) = \sqrt{(2c \ln \ln \ell)/ \ln \ell}$ into the proof of Theorem 3(b) in [5]. (We may not use Theorem 3(b) in [5] directly as stated as it needs $\nu(\ell)$ to be non-decreasing in ℓ.) $\square$

We use a result on graph discrepancy to prove Theorem 2. Given a graph $G = (V, E)$, the *discrepancy* of a set $X \subseteq V$ is defined as

$$D(X) := e(X) - \frac{1}{2}\binom{|X|}{2},$$

where $e(X)$ denotes the number of edges in the subgraph $G[X]$ induced by X. We use the following result of Erdős and Spencer [2, Ch. 7].

Lemma 4 (Theorem 7.1 of [2]). *Provided n is large enough and $t \in \mathbb{N}$ satisfies $\frac{1}{2}\log_2 n < t \le n$, then any graph $G = (V, E)$ of order n satisfies*

$$\max_{S \subseteq V, |S| \le t} |D(S)| \ge \frac{t^{3/2}}{10^3}\sqrt{\ln(5n/t)}.$$

Proof of Theorem 2. Let $G = (V, E)$ be any graph on at least $N = \lceil Ck \ln k \rceil$ vertices for a sufficiently large choice of C. We may assume that $k > \frac{1}{2}\log_2 N$ because otherwise G or $\overline{G}$ contains a clique of order k by the Erdős-Szekeres bound [3] on ordinary Ramsey numbers.

For any $X \subseteq V$ and $\nu > 0$, we define the following skew form of discrepancy:

$$D_\nu(X) := |D(X)| - \nu |X|^{3/2}.$$

We now construct a sequence $(H_0, H_1, \dots, H_t)$ of graphs as follows. Let H_0 be G or $\overline{G}$. At step $i+1$, we form H_{i+1} from $H_i = (V_i, E_i)$ by letting $X_i \subseteq V_i$ attain the maximum skew discrepancy D_ν and setting $V_{i+1} := V_i \setminus X_i$ and $H_{i+1} := H[V_{i+1}]$. We stop after step $t+1$ if $|V_{t+1}| < \frac{1}{2}N$. Let $I^+ \subseteq \{1, \dots, t\}$ be the set of indices i for which $D(X_i) > 0$. By symmetry, we may assume

$$\sum_{i \in I^+} |X_i| \ge \frac{1}{4}N. \tag{1}$$

Claim 1. *For any $i \in I^+$ and $x \in X_i$, $\deg_{H_i}(x) \ge \frac{1}{2}(|X_i| - 1) + \nu(|X_i| - 1)^{1/2}$.*

Proof. Write $|X_i| = n_i$. We are trivially done if $n_i = 1$, so assume $n_i \ge 2$. Suppose $x \in X_i$ has strictly smaller degree than claimed and set $X_i' := X_i \setminus \{x\}$. Then, since $i \in I^+$,

$$\begin{aligned} D_\nu(X_i') &\ge e(X_i') - \frac{1}{2}\binom{n_i - 1}{2} - \nu(n_i - 1)^{3/2} \\ &> e(X_i) - \frac{1}{2}\binom{n_i}{2} - \nu\sqrt{n_i - 1} - \nu(n_i - 1)^{3/2}. \end{aligned}$$

Note that $n_i^{3/2} > n_i^{1/2} + (n_i - 1)^{3/2}$, which by the above implies $D_\nu(X_i') > D_\nu(X_i)$, contradicting the maximality of $D_\nu(X_i)$. $\lozenge$

Claim 1 implies that we may assume for each $i \in I^+$ that $|X_i| \le k - 1$, or else we are done. This gives for any $i_1, \dots, i_4 \in I^+$ that

$$\left(\sum_{s=1}^{4} |X_{i_s}|\right)^{3/2} \le 8(k-1)^{3/2}. \tag{2}$$

Writing $I^+ = \{i_1, \dots, i_m\}$, we next show the following.

Claim 2. *For any $\ell \in \{1, \dots, m-3\}$, $D(X_{i_{\ell+3}}) \le \frac{5}{6}D(X_{i_\ell})$.*

Proof. For $X \subseteq V$, let us write $\nu(X) := \nu|X|^{3/2}$ so that $D_\nu(X) = |D(X)| - \nu(X)$. For $i_1, \ldots, i_r \in I^+$, we may write $X_{i_1,\ldots,i_r} := \bigcup_{s=1}^r X_{i_s}$. For disjoint $X, Y \subseteq V$, we define the *relative discrepancy* between X and Y to be

$$D(X,Y) := e(X,Y) - \frac{1}{2}|X||Y|,$$

where $e(X,Y)$ denotes the number of edges between X and Y.

Now let $i, j \in I^+$ with $i < j$. Then, by the maximality of $D_\nu(X_i)$, we have $D_\nu(X_i \cup X_j) \leq D_\nu(X_i)$, i.e.

$$|D(X_i) + D(X_i, X_j) + D(X_j)| - \nu(X_{i,j}) \leq |D(X_i)| - \nu(X_i) = D(X_i) - \nu(X_i),$$

and hence

$$D(X_j) \leq -D(X_i, X_j) + \nu(X_{i,j}). \tag{3}$$

Applying (3) (and the fact that $\nu(X_{i_{\ell+r}, i_{\ell+s}}) \leq \nu(\bigcup_{s=0}^3 X_{i_{\ell+s}})$ for any $r, s \in \{0,1,2,3\}$), we find that

$$D(X_{i_{\ell+1}}) + 2D(X_{i_{\ell+2}}) + 3D(X_{i_{\ell+3}}) \leq -\sum_{0 \leq r < s \leq 3} D(X_{i_{\ell+r}}, X_{i_{\ell+s}}) + 6\nu(\textstyle\bigcup_{s=0}^3 X_{i_{\ell+s}}). \tag{4}$$

Using $-D(\bigcup_{s=0}^3 X_{i_{\ell+s}}) - \nu(\bigcup_{s=0}^3 X_{i_{\ell+s}}) \leq D_\nu(\bigcup_{s=0}^3 X_{i_{\ell+s}}) \leq D_\nu(X_{i_\ell})$, we obtain

$$-\sum_{s=0}^{3} D(X_{i_{\ell+s}}) - \sum_{0 \leq r < s \leq 3} D(X_{i_{\ell+r}}, X_{i_{\ell+s}}) \leq D(X_{i_\ell}) + \nu(\textstyle\bigcup_{s=0}^3 X_{i_{\ell+s}}),$$

which combined with (4) implies that $D(X_{i_{\ell+2}}) + 2D(X_{i_{\ell+3}}) \leq 2D(X_{i_\ell}) + 7\nu(\bigcup_{s=0}^3 X_{i_{\ell+s}})$. From this, we obtain that

$$3D(X_{i_{\ell+3}}) \leq 2D(X_{i_\ell}) + 8\nu(\textstyle\bigcup_{s=0}^3 X_{i_{\ell+s}}), \tag{5}$$

where we have used the fact that $D(X_{i_{\ell+3}}) \leq D(X_{i_{\ell+2}}) + \nu(\bigcup_{s=0}^3 X_{i_{\ell+s}})$, which follows since $D_\nu(X_{i_{\ell+3}}) \leq D_\nu(X_{i_{\ell+2}})$. Using the fact that the graph H_{i_s} for any $s \in \{1, \ldots, m\}$ has at least $\frac{1}{2}N \geq \frac{C}{2}k \ln k$ vertices, it follows by Lemma 4 (using our assumption on k) that there exists a subset $Y_s \subseteq V_{i_s}$ of size at most k which satisfies

$$|D(Y_s)| \geq k^{3/2} \frac{\sqrt{\ln(C \ln k)}}{10^3}.$$

However, by our choice of X_{i_s}, we have

$$\begin{aligned} D(X_{i_s}) \geq D_\nu(X_{i_s}) \geq D_\nu(Y_s) &\geq |D(Y_s)| - \nu k^{3/2} \\ &\geq k^{3/2}\left(\frac{\sqrt{\ln(C \ln k)}}{10^3} - \nu\right) \geq 2\left(8\nu\left(\bigcup_{s=0}^{3} X_{i_{\ell+s}}\right)\right), \end{aligned}$$

by (2), provided C is sufficiently large. Therefore, from (5) we find that $3D(X_{i_{\ell+3}}) \leq 2D(X_{i_\ell}) + \frac{1}{2}D(X_{i_\ell})$, proving the claim. $\Diamond$

Claim 2 now implies that $(5/6)^{(m-1)/3} D(X_{i_1}) \geq D(X_{i_m}) \geq 1$ (assuming for simplicity $m \equiv 1 \pmod 3$), which then implies

$$m - 1 \leq \frac{3\ln(D(X_{i_1}))}{\ln(6/5)} \leq \frac{6}{\ln(6/5)} \ln(k-1).$$

By (1), we deduce that at least one of the m sets X_i with $i \in I^+$ satisfies

$$|X_i| \geq \frac{N \ln(6/5)}{25 \ln k}.$$

This last quantity is at least k by a choice of C sufficiently large, contradicting our assumption that $|X_i| \leq k-1$ for each $i \in I^+$. This completes the proof. $\square$

3 Subgraphs of high minimum degree via set-system discrepancy

In this section we prove, based on a well known discrepancy result of Spencer [8], that from a graph on $\ell = Ck$ vertices with minimum degree at least $\ell/2 + C'\sqrt{\ell}$ (with C' depending on C) we can select a subgraph on k vertices that has minimum degree at least $k/2$.

We start by defining the various standard notions of discrepancy that we need. Suppose $\mathcal{H} = \{A_1, \ldots, A_n\}$ where $A_i \subseteq V = [n]$. Let $\chi : V \to \{-1, 1\}$ be a colouring of V with the colours -1 and 1. For any $S \subseteq V$, we write $\chi(S) := \sum_{i \in S} \chi(i)$ and we define the *discrepancy* of $\mathcal{H}$ to be

$$\operatorname{disc}(\mathcal{H}) := \min_{\chi \in \{-1,1\}^V} \max_{S \in \mathcal{H}} \chi(S).$$

The result of Spencer [8] states that for any such $\mathcal{H}$ we have $\operatorname{disc}(\mathcal{H}) \leq 6\sqrt{\mathrm{n}}$.

For $X \subseteq V$, we define $\mathcal{H}|_X := \{A_1 \cap X, \ldots, A_n \cap X\}$. Then the *hereditary discrepancy* of $\mathcal{H}$ is defined by

$$\operatorname{herdisc}(\mathcal{H}) := \max_{\mathrm{X} \subseteq \mathrm{V}} \operatorname{disc}(\mathcal{H}|_{\mathrm{X}}).$$

The result of Spencer also immediately implies that $\operatorname{herdisc}(\mathcal{H}) \leq 6\sqrt{\mathrm{n}}$ for any $\mathcal{H}$.

Let A be the incidence matrix of $\mathcal{H}$, i.e. A is the $n \times n$ matrix given by

$$A_{ij} = \begin{cases} 1 & \text{if } j \in A_i, \\ 0 & \text{otherwise.} \end{cases}$$

Then we clearly have

$$\operatorname{disc}(\mathcal{H}) = \min_{\mathrm{x} \in \{-1,1\}^{\mathrm{V}}} \|\mathrm{Ax}\|_\infty = 2 \min_{\mathrm{x} \in \{0,1\}^{\mathrm{V}}} \left\| \mathrm{A}\left(\mathrm{x} - \frac{1}{2}\mathbb{1}\right) \right\|_\infty,$$

where $\mathbb{1}$ is the all 1 vector.

Now we define the *linear discrepancy* by

$$\operatorname{lindisc}(\mathcal{H}) := \max_{\mathrm{c} \in [0,1]^{\mathrm{V}}} \min_{\mathrm{x} \in \{0,1\}^{\mathrm{V}}} \|\mathrm{A}(\mathrm{x} - \mathrm{c})\|_\infty. \tag{6}$$

Note that here we are using $\{0, 1\}$-colourings again. Similarly, we define the hereditary linear discrepancy of $\mathcal{H}$ by

$$\operatorname{herlindisc}(\mathcal{H}) := \max_{\mathrm{X} \subseteq \mathrm{V}} \operatorname{lindisc}(\mathcal{H}|_{\mathrm{X}}).$$

A result of Lovász, Spencer, and Vestergombi [6] states that herlindisc($\mathcal{H}$) $\leq$ herdisc($\mathcal{H}$). (Note that the factor of 2 from [6] is missing to adjust for the slightly different definition we are using.) Combining with Spencer's result, we have

$$\mathrm{lindisc}(\mathcal{H}) \leq \mathrm{herlindisc}(\mathcal{H}) \leq \mathrm{herdisc}(\mathcal{H}) \leq 6\sqrt{\mathrm{n}}.$$

If we set c to be the all p vector (for some $p \in [0,1]$) in (6), we obtain the following result.

Lemma 5. *Let $A_1, \ldots, A_n \subseteq V = [n]$ and $p \in [0,1]$. Then there exists $Y \subseteq V$ such that, for all $i \in [n]$,*

$$||A_i \cap Y| - p|A_i|| \leq 6\sqrt{n}.$$

We use the previous lemma to prove the following result.

Lemma 6. *Suppose $G = (V, E)$ is a graph with $\ell = Pk$ vertices for some $P > 1$ and k a positive integer, and suppose*

$$\delta(G) \geq \frac{1}{2}\ell + \eta\sqrt{\ell}$$

for some $\eta > 0$. Then G has an induced subgraph H on k vertices with minimum degree

$$\delta(H) \geq \frac{1}{2}k + \left(\frac{\eta}{\sqrt{P}} - 19\sqrt{P}\right)\sqrt{k}.$$

Proof. Write $V = \{v_1, \ldots, v_\ell\}$, let $A_0 = V$ and for each $i \in [\ell]$ let $A_i \subseteq V$ be the neighbourhood of v_i in G. We apply Lemma 5 to the sets $A_0, \ldots, A_{\ell-1}$ with $p = (k + 1 + 6\sqrt{\ell})/\ell$. (Note that if $p > 1$ then with a simple calculation it is easy to see we can obtain the desired graph H simply by deleting any $\ell - k$ vertices from G.) Thus there exists $Y \subseteq V$ satisfying

$$||A_i \cap Y| - p|A_i|| \leq 6\sqrt{\ell}$$

for all $i \in \{0, \ldots, \ell - 1\}$. Applying this for $i = 0$ and noting $A_0 \cap Y = Y$ gives

$$k + 1 = p|A_0| - 6\sqrt{\ell} \leq |Y| \leq p|A_0| + 6\sqrt{\ell} = k + 1 + 12\sqrt{Pk}$$

and applying it for $i \in [\ell - 1]$ gives

$$\begin{aligned} |A_i \cap Y| \geq p|A_i| - 6\sqrt{\ell} \geq \frac{k}{\ell}\left(\frac{1}{2}\ell + \eta\sqrt{\ell}\right) - 6\sqrt{\ell} &= \frac{1}{2}k + \eta\frac{k}{\sqrt{\ell}} - 6\sqrt{\ell} \\ &= \frac{1}{2}k + \left(\frac{\eta}{\sqrt{P}} - 6\sqrt{P}\right)\sqrt{k}. \end{aligned}$$

Thus Y has between $k + 1$ and $k + 1 + 12\sqrt{P}\sqrt{k}$ vertices. Let Z be an arbitrary subset of $Y \setminus \{v_\ell\}$ of size k and let $H = G[Z]$. Then since we have removed at most $12\sqrt{Pk} + 1 \leq 13\sqrt{Pk}$ vertices from Y to obtain Z, we have for each $i \in [\ell - 1]$ that

$$|A_i \cap Z| \geq \frac{1}{2}k + \left(\frac{\eta}{\sqrt{P}} - 19\sqrt{P}\right)\sqrt{k}.$$

In particular this means

$$\delta(H) \geq \frac{1}{2}k + \left(\frac{\eta}{\sqrt{P}} - 19\sqrt{P}\right)\sqrt{k},$$

as desired. $\square$

4 Proof of Theorem 1

To prove the theorem, we use as a subroutine the following algorithm, which is inspired by the greedy algorithm for max-cut or min-bisection.

Lemma 7. *Let $G = (V, E)$ be a graph of order n with $\delta(G) \geq \frac{1}{2}(n-1) + t$ for some number t. Let $\alpha \in [0, 1]$ and let $a, b \in \mathbb{N}$ such that $a + b = n$. Then either there exists $A \subseteq V$ of size a such that $\delta(G[A]) \geq \frac{1}{2}a - 1 + \alpha t$, or there exists $B \subseteq V$ of size b such that $\delta(G[B]) \geq \frac{1}{2}b - 1 + (1-\alpha)t$.*

Proof. Take any $A \subseteq V$ of size a and let $B := V \setminus A$. If there exists $x \in A$ with $\deg_A(x) < \frac{1}{2}a - 1 + \alpha t$ and $y \in B$ with $\deg_B(y) < \frac{1}{2}b - 1 + (1-\alpha)t$, then move x to B and y to A, i.e. *swap* x and y. Note that when there is no such pair of vertices x, y we are done. We just need to prove that, if we keep iterating, then this procedure must stop at some point.

Consider the number of edges in $G[A]$ before and after we swap x and y. The number of edges in $G[A]$ increases by at least

$$\deg_A(y) - \deg_A(x) - 1 \geq \delta(G) - \deg_B(y) - \deg_A(x) - 1 \geq 1/2,$$

(where we subtracted 1 in case x and y are adjacent). This shows that we cannot continue to swap pairs indefinitely. $\square$

At last we are ready to prove the main result. In fact, we prove something stronger.

Theorem 8. *There exist constants $D, D' > 0$ such that for $k \geq 2$ and any graph G on $Dk \ln k$ vertices, G or its complement $\overline{G}$ has an induced subgraph on k vertices with minimum degree at least $\frac{1}{2}(k-1) + D'\sqrt{(k-1)/\ln k}$.*

Proof. Set $\nu = 160$, $C = C(\nu)$ as defined according to Theorem 2, and $D := 4C$. Also set $D' := 1/\sqrt{D}$.

By Theorem 2, since $C \cdot 2k\ln(2k) \leq 4Ck\ln k = Dk\ln k \leq |V(G)|$, we find G or $\overline{G}$ has an induced subgraph H on $\ell \geq 2k$ vertices with $\delta(H) \geq \frac{1}{2}(\ell-1) + \nu\sqrt{\ell-1}$.

Let $x = \ell \bmod k$ (so $x \in \{0, \dots, k-1\}$). We can now apply Lemma 7 to H with $a = k + x$, $b = \ell - k - x$, $t = \nu\sqrt{\ell-1}$ and $\alpha = 1/2$. Suppose this gives us a subset $A \subseteq V(H)$ of size a such that

$$\delta(H[A]) \geq \frac{1}{2}a - 1 + \frac{1}{2}\nu\sqrt{\ell-1} \geq \frac{1}{2}a + \frac{1}{4}\nu\sqrt{\ell} \geq \frac{1}{2}a + \frac{1}{4}\nu\sqrt{a}.$$

Then $k \leq a < 2k$ and, so applying Lemma 6 (with $P = a/k \in [1,2]$ and $\eta = \nu/4 = 40$) yields a subset $A' \subseteq A$ of size k such that

$$\delta(H[A']) \geq \frac{1}{2}k + \left(\frac{40}{\sqrt{P}} - 19\sqrt{P}\right)\sqrt{k} \geq \frac{1}{2}k + \left(\frac{40}{\sqrt{2}} - 19\sqrt{2}\right)\sqrt{k} \geq \frac{1}{2}k + \sqrt{2k},$$

which is more than required. In case Lemma 7 does not produce such a set A, it gives instead a subset B of size $b = \ell - k - x \equiv 0 \pmod{k}$ such that $\delta(H[B]) \geq \frac{1}{2}(b-1) + \frac{1}{2}\nu\sqrt{\ell-1} - \frac{1}{2}$. We iteratively apply Lemma 7 to $H[B]$ in a binary search to find a desired induced subgraph as follows.

Set $G_0 = H[B]$. Let $\ell_0 := |V(G_0)| = b$ (so that $k \leq \ell_0 \leq Dk\ln 2k$ and $\ell_0 \equiv 0 \pmod{k}$) and set $t_0 := \frac{1}{2}\nu\sqrt{\ell-1} - \frac{1}{2} \geq \frac{1}{2}\nu\sqrt{\ell_0-1} - \frac{1}{2}$ (so that $\delta(G_0) \geq \frac{1}{2}(\ell_0-1) + t_0$). Suppose that G_i is given, where G_i has ℓ_i vertices with $\ell_i \equiv 0 \pmod{k}$ and $\delta(G_i) \geq \frac{1}{2}(\ell_i-1) + t_i$ for some number t_i. Set $a_i = \lfloor \ell_i/2k \rfloor k$ and $b_i = \lceil \ell_i/2k \rceil k$ so that $a_i + b_i = \ell_i$ and $a_i \equiv b_i \equiv 0$

(mod k). Apply Lemma 7 with $G = G_i$, $a = a_i$, $b = b_i$, $t = t_i$, and $\alpha = \frac{1}{2}$. Then we either obtain a set of vertices A_i of size a_i such that $\delta(G_i[A_i]) \geq \frac{1}{2}a_i - 1 + \frac{1}{2}t_i$, in which case we set $G_{i+1} := G_i[A_i] = H[A_i]$, or we obtain a set of vertices B_i of size b_i such that $\delta(G_i[B_i]) \geq \frac{1}{2}b_i - 1 + \frac{1}{2}t_i$, in which case we set $G_{i+1} := G_i[B_i] = H[B_i]$. Now set $\ell_{i+1} = |V(G_{i+1})|$ and note that $\ell_{i+1} \equiv 0 \pmod{k}$ and $\delta(G_{i+1}) \geq \frac{1}{2}(\ell_{i+1} - 1) + t_{i+1}$, where $t_{i+1} = \frac{1}{2}(t_i - 1)$. Note also that $\ell_{i+1}/k \leq \lceil \ell_i / 2k \rceil$.

In this way we obtain subgraphs $G_0, G_1, \ldots$ of $G_0 = H[B]$ and we see from the recursion for ℓ_i above that if $\ell_i > k$ then $\ell_{i+1} < \ell_i$. Thus there exists some j such that $\ell_j = k$ (since $\ell_i \equiv 0 \pmod{k}$ for all i) and an easy computation shows we can assume that $j \leq \log_2(\ell_0/k) + 1$. The recursion for t_i implies that $t_i \geq t_0 2^{-i} - 1$ so that

$$t_j \geq \frac{t_0 k}{2\ell_0} - 1 \geq \frac{\nu(\sqrt{\ell_0 - 1} - 1)k}{4\ell_0} \geq \frac{k}{\sqrt{\ell_0}} \geq \frac{\sqrt{k}}{\sqrt{D \ln k}} = D' \sqrt{\frac{k}{\ln k}}$$

(where we used that $t_0 \geq \frac{1}{2}\nu\sqrt{\ell_0 - 1} - \frac{1}{2}$, that $\ell_0 \geq k \geq 2$ with $\nu = 160$, and that $\ell_0 \leq Dk \ln k$). Thus G_j has k vertices and minimum degree at least $\frac{1}{2}(k-1) + D'\sqrt{(k-1)/\ln k}$ and is an induced subgraph of $H[B]$ and hence of G or $\overline{G}$. □

5 Concluding remarks

It is tempting to try using the greedy subroutine (Lemma 7) in a binary search on the output of Theorem 3(a) of [5], but since we cannot control the order of this output graph, the search might require $O(\ln k)$ steps, which would destroy the minimum degree bounds.

Determination of the second-order term in the minimum degree threshold for polynomial to super-polynomial growth of the fixed quasi-Ramsey numbers is an open problem. (The corresponding term for the variable quasi-Ramsey numbers was determined in [5].) To pose the problem concretely, we recall notation of Erdős and Pach. For $c \in [0,1]$ and $k \in \mathbb{N}$, let $R^*_c(k)$ be the least number n such that for any graph $G = (V, E)$ on at least n vertices, there exists $S \subseteq V$ with $|S| = k$ such that either $\delta(G[S]) \geq c(k-1)$ or $\delta(\overline{G}[S]) \geq c(k-1)$. Now consider $c = \frac{1}{2} + \varepsilon$ where $\varepsilon = \varepsilon(k)$ is a function of the size k of the subset sought. By Theorem 8 if $\varepsilon(k) = O(\sqrt{1/(k-1)\ln k})$ then $R^*_c(k)$ is polynomial in k, and by Proposition 3 if $\varepsilon(k) = \omega(\sqrt{\ln\ln k/(k-1)})$ then $R^*_c(k)$ is superpolynomial in k. Hence the choice of ε for which we find a transition between polynomial and super-polynomial growth in k of $R^*_c(k)$ is essentially determined to within a $\sqrt{\ln k \ln\ln k}$ factor of $\sqrt{1/(k-1)}$. What is it precisely?

Last, we remark that, in the above notation, our main result is that $R^*_{1/2}(k) \leq Ck\ln k$ for some $C > 0$, while Erdős and Pach showed that $R^*_{1/2}(k) \geq C'k\ln k/\ln\ln k$ for some $C' > 0$. They also asked if $R^*_{1/2}(k) \geq C'k\ln k$ for some $C' > 0$. This question remains open.

Acknowledgement

We thank Noga Alon for stimulating discussions during ICGT 2014 in Grenoble. We are grateful to Joel Spencer for helpful comments about linear and hereditary discrepancy.

References

[1] P. Erdős and J. Pach. On a quasi-Ramsey problem. *J. Graph Theory*, 7(1):137–147, 1983.

[2] P. Erdős and J. Spencer. *Probabilistic Methods in Combinatorics*. Probability and Mathematical Statistics, A Series of Monographs and Textbooks. Academic Press, New York and London, 1974.

[3] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.

[4] R. L. Graham, B. L. Rothschild, and J. H. Spencer. *Ramsey theory*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., New York, second edition, 1990. A Wiley-Interscience Publication.

[5] R. J. Kang, J. Pach, V. Patel, and G. Regts. A precise threshold for quasi-Ramsey numbers. *SIAM J. Discrete Math.*, 29:1670–1682, 2015.

[6] L. Lovász, J. Spencer, and K. Vesztergombi. Discrepancy of set-systems and matrices. *European J. Combin.*, 7(2):151–160, 1986.

[7] H. J. Prömel. *Ramsey theory for discrete structures*. Springer, Cham, 2013. With a foreword by Angelika Steger.

[8] J. Spencer. Six standard deviations suffice. *Trans. Amer. Math. Soc.*, 289(2):679–706, 1985.